# The Relationship Between Reward and Probability: Evidence that Exploration may be Intrinsically Rewarding

Rosie Clark and Iain D. Gilchrist

University of Bristol

School of Psychological Science,
University of Bristol
12A Priory Road
Bristol,
BS8 1TU,
U.K.

Corresponding Author: Dr Rosie Clark
r.clark@bristol.ac.uk
+441173313303

Iain D. Gilchrist
i.d.gilchrist@bristol.ac.uk
+441179289004

Abstract

The probability of an event occurring and the reward associated with the event can both modulate behaviour. Response times are decreased to stimuli that are either more rewarding or more likely. These two factors can be combined to give an *Expected Value* (EV) associated with the event (i.e. probability of the event x reward magnitude). In four experiments we investigate the effect of reward and probability on both saccadic and manual responses. When tested separately we find evidence for both a reward and probability effect across response types. When manipulations of reward magnitude and probability of the event were combined, the probability modulations dominated and these data were not well accounted for by the EV. However, a post-hoc model that included an additional intrinsic reward associated with responding provided an excellent account for the data. We argue that reward consists of both an explicit and intrinsic component. In our task, the saccadic and manual responses are linked to the information provided by the targets and the goals of the task, and successful completion of these is in itself rewarding. As a result targets associated with a higher probability of being presented have a higher intrinsic reward.

Keywords

probability, reward, saccades, expected value

## Introduction

Human behaviour is shaped by both the properties of the environment and of the individual. In any exploration task one property of the environment that is important is the likelihood of the target occurring in a particular location. Equally a property of the individual that will shape behaviour is the particular value that they place on the target. Take for example a child searching for a cuddly toy to take to bed to help them drop off to sleep. She may know that it is quite likely that her favourite elephant was left behind at her grandparents and so is unlikely to be in the house and at the same time know that her less preferred squirrel is almost certainly in the kitchen. How is the child meant to combine the probability information with the reward value information associated with her two toys to shape her search behaviour?

One way in which these factors can combine is by Expected Value (EV: the product of the probability of rewarded event and the magnitude of reward associated with it). In the economic literature, the expected value of future rewards is the sum of possible reward magnitudes where each is weighted by its probability (Tobler, Fiorillo, & Schultz, 2005). When considering rewards that occupy spatial locations, like the example above, there are two distinct ways in which probability can be defined. In the first, 'probability' is the probability of getting a reward from a given object at a location. In the second the 'probability' is the likelihood of an object being in a given location, but the reward associated with that object is fixed. Is it the second of these cases that is illustrated by the example of the child looking for their toy and is the focus of the current paper. Such condition haves been investigated by Milstein and Dorris (2007), who manipulated target reward probability in dual target trials by varying the probability of left vs right target presentation and manipulated the reward magnitude associated with each target. They found that EV predicted saccadic behaviour better than the probability of a rewarded event or reward magnitude alone.

Allowing behaviour to be shaped by EV has an evolutionary advantage as it means that the individual will maximise the intake of reward over time (Milstein & Dorris, 2007). This behavioural data is supported by neurophysiological studies showing that dopamine neurons, known to play an important role in reward processing, encode a combination of reward magnitude and probability (e.g. Expected Value) rather than distinguishing between the two (Tobler, Fiorillo, & Schultz, 2005). Understanding of this system has partly arisen from neurological and psychiatric disorders involving changes in dopaminergic levels (e.g. schizophrenia, Parkinson's disease, drug addiction) where processing of reward is dysfunctional (Dreher & Tremblay, 2009).

Supporting this, many behavioural studies in non-human primates have shown that saccadic responses have shorter latencies to targets that are associated with reward than to those that are not (Kawagoe, Takikawa, & Hikosaka, 1998; Takikawa, Kawagoe, Itoh, Nakahara, & Hikosaka, 2002). Additionally, Milstein and Dorris (2007, 2011) have shown that it is the relative value of a stimulus that affects choice and saccadic reaction times, rather than the global magnitude of reward available in a trial. They found that saccadic preparation is spatially allocated based on the relative value of potential targets. This is also supported by the economic literature: prospect theory posits that value (or utility) of an action is relative to the other available options (Kahneman & Tversky, 1979). Allocating time and resources towards more profitable options relative to others may be more adaptive than being globally motivated by the prospect of reward. Several studies have also shown that these reward effects persist even after reward is no longer associated with these stimuli (Dunne, Ellison, & Smith, 2015; Stankevich & Geng, 2015; Theeuwes & Belopolsky, 2012).

The probability of the occurrence of stimuli in different locations has long been known to affect our response (Kaufman & Levy, 1966; Lamb & Kaufman, 1965; Mowbray, 1964) and the probability effect is particularly well established for a saccadic response (Basso &

Wurtz, 1997; Carpenter & Williams, 1995; Dorris & Munoz, 1998; Jóhannesson, Haraldsson & Kristjánsson, 2013; Koval, Ford & Everling, 2004; Liu et al., 2010, 2011; Noorani & Carpenter, 2013). For example, it has been shown that for two target locations, decreasing the relative probability of target presentation in a location increases the saccadic reaction time (Carpenter & Williams, 1995).

Only a few studies (e.g. Milstein and Dorris, 2007, 2011) have examined the relationship between the processing of reward and probability information, and the mechanisms that could govern these effects. One common mechanism for the effect of both factors on response times could be attention. Theeuwes and Belopolsky (2012) demonstrated a link between increases in the associating reward with a target and increased attentional capture. Studies of the effect of target probability have used a pro- and anti-saccade task alongside a probability manipulation (Koval et al., 2004; Liu et al., 2010) and proposing that the higher location probabilities demand more attention, based on the finding that the anti-saccade cost (longer latencies and more errors) was only present in anti-saccades trials away from the presented high probability target location (Liu et al., 2010). This attentional framework explains the increased difficulty in inhibiting prosaccades to the high probability location during antisaccades to the opposite location and provides an explanation as to why pro-saccades to lower probability locations take longer to be executed (Liu et al., 2010).

One alternative way to conceptualise the probability of a stimulus is as *expectation*, which could guide attention towards a likely location of a relevant item (Summerfield & Egner, 2009). However, the similarities between attention and expectation are not reflected in the neural activity in visual regions representing the stimulus, and some neurophysiological research has shown that reward (defined as "motivation") and attention contribute independently to influence the responses (Bendiksby & Platt, 2006; Summerfield & Egner, 2009). One way in which reward and probability may contribute to attention (but not directly

change it) in a similar way is by increasing the salience of the target that is associated with high reward value or probability. Many studies in the reward literature have linked reward associations with increasing salience, which is defined as the physical, "bottom-up" distinctiveness of an object relative to other objects in the environment (Fecteau & Munoz, 2006).

Supporting these behavioural similarities between probability and reward processing, Nakahara, Nakamura and Hikosaka (2006) demonstrated that prior knowledge about the reward associated with a stimulus leads to a bias in the excitability of SC neurons and neural activity in the superior colliculus (SC). These effects preceded target selection and were found to increase as the probability (of any rewarded event) increased (Basso & Wurtz, 1997; Dorris & Munoz, 1998; Liu et al., 2011). In addition, a study separately investigating both prior knowledge of reward (probability of receiving reward from given response rather than differing magnitude) or probability of target location found that these factors shape perception and action in parallel and suggests that a shared sensory weight amplifies perceptual experience while biasing motor action driven by attention and EV (Liston & Stone, 2008). These results suggest that the effects of reward and probability on decision-making processes across different response modalities may be governed by the same system or overlap in some way.

The majority of previous human and primate studies on manipulations of reward and probability have focused on one response modality. The premotor theory of attention postulates that processes involved in spatial attention and selection of motor responses share a common neural substrate (Eimer, Van Velzen, Gherri, & Press, 2006; Rizzolatti, Riggio, & Sheliga, 1994). This theory suggests that attentional shifts arise as a by-product of activity within the oculomotor system. Within this model any effects on attention should be common and correlated across both oculomotor and manual response preparation (Eimer et al., 2006;

Rizzolatti et al., 1994). If common attentional modulation underpins the effects of both reward and probability on behaviour then, in turn, this suggests that both reward and probability should affect both oculomotor and manual response preparation in a similar and correlated way. Decision-making models for visuomotor behaviour generally propose that a visual event produces an internal response, rising with time until a decision threshold is reached and a motor response is initiated (Bompas & Sumner, 2008). Within the pre-motor theory of attention, a single such process would be used for all types of motor responses and the same decision threshold would apply to all responses: response time would then be the sum of stimulus processing time and motor latency where the latter is the only source of difference between response types (Bompas & Sumner, 2008). A number of studies have shown support for this theory, for example Taylor, Carpenter and Anderson (2006) found similar parameters in manual and saccadic response times for visual processing of contrast, supporting a common target detection stage preceding each type of reaction. Research by Bompas and Sumner (2008) partially supports these differences between processes, but specifically investigating manual and saccadic reaction times. They found that the difference in reaction times to S-cone (blue cone) and luminance signals was larger for saccade latencies than for manual responses (Bompas & Sumner, 2008). This could suggest that saccadic responses can take better advantage of fast signals when they are available (Bompas & Sumner, 2008). This result is not compatible with the suggestion that the same target detection stage is used for manual and saccadic responses (Taylor et al., 2006). An alternative, but less parsimonious model, would propose a common collector stage where signals across all pathways are brought together but have different decision thresholds for different responses (Bompas & Sumner, 2008). However, Bompas and Sumner (2008) found no evidence for a correlation between saccadic and manual reaction times; a correlation would be expected if they both relied on the same combination of signals. Together these

results suggest a more complex explanation beyond the premotor theory of attention for the similarities and differences between manual and saccadic responses, but do provide a strong rational for further investigation in this area.

Here we report a series of four experiments, the first three of which concurrently investigate saccadic and manual responses. Across all four experiments, participants were required to look at a single target presented on the left or right of a screen. The target contained a capital letter T and the participants had to respond with a different manual response if the T was the correct orientation or inverted. In Experiment 1 we manipulated the probability of a target being presented on the left or the right-hand side. In Experiment 2 we investigated the absolute and relative reward associated with two targets. In Experiment 3 we combined these two manipulations to examine whether these factors combine in a simple manner and in Experiment 4 we examined the two manipulations separately within participants.

## General Method

### *Procedure*

*Figure 1.*

The sequence of events in a single example trial is illustrated in Figure 1. The task for the participants was to respond to the "T" stimuli that appeared to the left or right of a central fixation point. The response required was to indicate the orientation of a letter with a manual button press. In Experiments 1 to 3, if the letter was on the left-hand side of the display the response was made with the left hand and if the letter was on the right-hand side the response was made with the right hand. Each block in each of the four experiments was proceeded by a 9-point calibration procedure to allow accurate eye tracking.

All studies reported here were approved by The Faculty of Science Human Research Ethics Committee at the University of Bristol.

### *Stimuli*

All stimuli and the fixation square were white (16.4 cd/$m^2$) and displayed on a grey background (10.4 cd/$m^2$). A trial commenced with a centrally presented fixation square (1.65$^o$ x 1.65$^o$; 0.18$^o$ thick) which was presented for 1.5s. This was immediately followed by the circular target with a T or inverted T in the centre. The circle had a diameter of 1.85$^o$ and a line thickness of 0.18$^o$. The letter T subtended 0.3$^o$, which is a size that Körner and Gilchrist (2007) have shown is small enough not to be recognised reliably above chance when fixation was 3$^o$ away from the stimuli. This ensured that participants had to make an accurate target directed saccadic eye movement towards the target to complete the task. The target was presented on the left or right in a varying number of trials across all phases of the four experiments at 6$^o$ eccentricity.

The pre-test block was included in all four experiments to set an individual criterion for the time-out for the testing phase. In the practice and pre-test phases the fixation square contained an X in red (15.8 cd/$m^2$; font size 17). After a successful response to the letter T, an X was presented in the target in a golden yellow colour (19.8 cd/$m^2$). These colours were chosen to match the reward stimuli colours in the later Experiments (2 to 4). The distribution of manual response times from the pre-test block of the experiment was used to calculate a $70^{th}$ percentile of each participant's reaction time distribution. Without informing the participants, their individual $70^{th}$ percentile values were used as the length of time the target and letter T would be visible in testing phase trials, after which they would receive the time-out notice. This ensured that participants were motivated to respond as quickly as possible.

During the testing phase of all four experiments in this paper, participants were clearly informed before each block of the probabilities (e.g. 'the probability of the target appearing on the right is 75%') and reward magnitudes (e.g. '£2') associated with each target.

*Apparatus*

All the experiments were controlled by Psychophysics Toolbox (Brainard, 1997) running within Matlab on Windows 7. The display was 17" running at 75Hz and with a resolution of 1600 x 1200 pixels, and the viewing distance was 57cm.

Movements of the right eye were recorded at a sampling rate of 1000Hz by the Eyelink II (SR Research, Canada) which has a typical operating spatial resolution of 0.5º. The participants were provided with a chin and forehead rest to minimise head movements. In experiments 1 to 3, manual responses were recorded via the keyboard (numeric right-hand section) – key 4 and 1 for the left hand and key 6 and 3 for the right-hand responses.

## Experiment 1: Probability

### Methods

#### *Procedure*

The layout of the task was as described in the general method. Experiment 1 consisted of a practice phase (10 trials), a pre-test phase (48 trials), and then 6 blocks (48 trials each) of the testing phase. The testing phase consisted of three block types. In the first, the target was equally likely to be on the right or the left-hand side. In the second, the target was presented on one side with probability 0.75 and the other side with probability 0.25. In the third, probabilities were then flipped. The three probability block types were split into two separate consecutive blocks of 48 trials. The probabilities were pseudo-random being fixed at 50%, 75% or 25% of the total number of trials in the block. The order of these fixed trials were then randomised for the length of the block.

#### *Stimuli*

All the stimuli were as described in the general method. In addition, during the pre-test phase the fixation square contained a plus sign as a fixation point, which was presented in red (15.8 cd/$m^2$; font size 17). No reward value was ever shown.

The participants also received feedback if they were too slow, or made an incorrect response. A message appeared in the centre of the screen reading “Wrong!” or “Too slow!” (2.5$^o$ above fixation and white). After the response to the letter T, the empty target location marker (or error message) stayed on the screen for 1.5 seconds.

*Participants*

Eighteen participants (8 male) were recruited from the student population of the University of Bristol (approximate age range 18-25). All had normal or corrected-to-normal vision. Participants were reimbursed £7 for their time and given no other additional reward incentive.

*Design*

There was one within participant repeated measure binary factor: probability block type (75/25, 50/50, 25/75). Combining data from the right and left sides lead to three conditions: 75% probability, 50% probability and 25% probability.

The orders of conditions were counterbalanced using a latin square design. The dependent measures were the manual response time and saccade latency.

*Manual Response Analysis*

A total of 5184 trials were recorded (288 x 18 participants). The total number of errors in manual responses (time-out/response errors) made was 550, leaving 4634 trials to be analysed. The range of the total percentage of time-out/response errors of each participant's responses was 4.17% to 21.9%. Each participant's median responses were analysed and this convention was maintained across the experiments reported in this paper.

*Saccadic Response Analysis*

A general exclusion criterion was applied to initial fixation starting point: excluding trials where the initial fixation was greater than 3 degrees either side of the centre of the fixation box (21.3% excluded, range across participants: 3.13-36.1%). This was a liberal criterion as there was no specific fixation instruction to participants in the experiment. An additional criteria was applied so that all saccades initiated after the participant's 'time-out'

were discarded. Two participant's data had over 40% of trials excluded at this point, and were therefore removed from the rest of the saccadic analysis. There was no spatial exclusion criterion on the landing point of the first saccade, but saccades toward the incorrect side were discarded as errors (3.1% excluded, range across participants: 0-7.29%). The remaining trials had a mean first saccade amplitude of 6.0 degrees (range across participants: 5.5 – 6.4 degrees). This shows that generally all participants were making a large hypo-metric orienting saccade for their first saccade.

Of the saccades directed to the incorrect side, 91.5 % had a latency of less than 90ms and these are assumed to be anticipatory. In addition to reflect this anticipatory criteria all saccades with reaction times less than 90ms in the correct direction were excluded (4.8%, range across participants: 0.69-12.2%). Some saccades within the incorrect saccades and anticipatory saccades categories had negative response times, due to the saccade being on-going during the target onset. These saccades were included in the anticipatory analysis as they still reflect an anticipation of the target appearing on a particular side. Six of the anticipatory saccades were removed before analysing the frequencies, due to either errors in the eye tracker, or saccades disrupted by blinks. After removing the invalid, anticipatory and error saccades, the total number of analysed saccades was 3260 (70.8%). Each participant's median responses were analysed.

## Results

### *Manual Responses*

*Table 1*.

*Figure 2.*

The manual results are plotted in Figure 2. A significant effect of target probability on manual response times was seen, with a 14 ms (95% CI: 6.31 – 21.8ms) increase in response time from 0.5 and 0.25 target probability, and a 21ms (95% CI: 10.9 – 30.5ms) decrease in response times between 0.5 and 0.75 probability (Repeated Measures ANOVA: [$F(2,17) = 26.65$, $p < 0.01$]; Effect size (partial ETA squared) = 0.61). Additionally, there was evidence that this effect had a significantly linear trend [$F(1,17) = 35.76$, $p < 0.001$].

Error rates are reported in Table 1. There was a significant effect of error type (time-out/response errors) (Repeated Measures ANOVA: [$F(1,17) = 8.2$, $p = 0.01$]; Effect size (partial ETA squared) = 0.33) and target probability (Repeated Measures ANOVA: [$F(2,17) = 4.55$, $p = 0.02$]; Effect size (partial ETA squared) = 0.21). There were slightly more time-out errors than response errors recorded, and the frequency of errors increased a small amount as the probability decreased.

*Saccadic Responses*

In the saccadic data there was a significant linear trend between the mean response times across the three spatial probabilities (Repeated Measures ANOVA within-subjects contrast: [$F(1, 15) = 9.84$, $p = 0.007$]; Effect size (Partial ETA squared) = 0.4) (Figure 2). No significant difference was found between the percentages of saccades that were anticipatory across the spatial probabilities. This was not present in the anticipatory saccades in the incorrect direction (Repeated Measures ANOVA: [$F(2, 15) = 2.08$ , $p = 0.14$]; Effect size (Partial ETA squared) = 0.58), or in the correct direction anticipatory saccades [$F(2, 15) = 2.2$, $p = 0.13$]

The relationship between the manual and saccadic responses within participants was investigated by calculating the size of the probability effect between 0.75 and 0.25 probability for each participant. The two participants excluded just from the saccadic analysis were excluded from these calculations. There was no significant correlation between the two response modalities (correlation coefficient = 0.11, N.S. p = 0.65).

## Experiment 1: Discussion

Our results support the existing literature showing that response times are increased as the probability is decreased, and conversely decreased as the probability is increased (Carpenter & Williams, 1995). However, the probability effect is clearest in the manual responses, compared to the saccadic responses. This supports the finding that within the saccadic system this effect is not always present or robust (Jóhannesson et al., 2013). One reason we might not have seen a strong saccadic effect is the length of the inter-stimulus interval (ISI; time from onset of fixation until onset of the target). The ISI was relatively long at a constant of 1.5 seconds. If we compare this ISI duration with the existing literature, Antonaides et al. (2014) replicated Carpenter and Williams' (1995) probability effect in healthy older adults using the same method with an ISI of between 0.5 and 1 second. In contrast, in a similar paradigm Jóhannesson et al. (2013) found no effect of probability (0.75/0.5/0.25) on saccadic response times: the ISI they used was 0.6 to 1.6 seconds. It could be that manipulations of target probability will only strongly modulate saccadic response times when the ISI is short (no longer than 1 second).

There is some discussion in the literature about the underlying mechanisms that lead to these probability effects. Two mechanisms have been proposed (see Walthew and Gilchrist,

2006). In the first, the effect of probability arises from the long-term learning of the locations that are more frequent. In the second, the effects are a result of inter-trial repetition. The argument here is that higher probability locations are more likely to be proceeded by a target at the location in the previous trial and it is these immediate repetitions that drive the probability effects. Under some circumstances it would appear that these local mechanisms alone support probability effects (Walthew and Gilchrist, 2006) although not always and this may depend on the exact experimental set-up (Druker & Anderson, 2010; Jones & Kaschak 2012). To investigate this is the current data we analysed the manual response time data as the effects were more robust for this response type. Repetitions of the previous location lead to a 15 ms decrease in reaction time (95% CI: 5 – 25ms) but the effect of probability was still present, and of a similar magnitude, when we compared high and low probability trials that were proceeded by the opposite type of trial: 31ms difference (95% CI: 15 – 46 ms) and trials that were proceeded by the same type of trial: 27ms difference (95% CI: 8 – 45 ms) . This suggests that the probability effect in this particular set-up is driven by a long-term learning mechanism rather than repetition, despite repetition effects being present.[1]

## Experiment 2: Lottery controlled reward

In Experiment 2 we investigated the facilitation of saccadic and manual response times by increasing reward value. We employed a random lottery incentive reward system, which means that one trial is selected at random from the whole experiment and the participant

[1] The means in this analysis are slightly different from the main results due to response errors (but not time-out errors) being included in this calculation. This is done to maximise the number of trials we have for this analysis particularly for the low probability trials.

receives the reward from that trial (Cubitt, Starmer, & Sugden, 1998). The validity of this system in economic experiments has been tested and shown to have no significant difference on behaviour in comparison to a design where participants are rewarded on every trial (Cubitt et al., 1998). This addresses the problem of the salience of rewards, as large monetary reward values can be used without having to rely on a points-based rewards system that may not provide enough of an explicit link to monetary rewards.

## Methods

### *Procedure*

The task layout was as described in the general method and similar to Experiment 1. The experiment consisted of a practice phase (10 trials), a pre-test phase (48 trials), and then 4 blocks (72 trials each) of the testing phase. In the testing phase correct responses on each trial were associated with a reward. The reward value was displayed on each trial at the location of the target after the response was given. Participants were informed that a trial from the testing phase would be selected at random at the end of the experiment and that they would receive that reward. If the manual response was too slow or incorrect then a message was displayed after the trial. Participants were informed at the outset of the experiment that they were to receive no reward if one of these error trials was selected at the end of the experiment. All participants were explicitly told the reward values associated with the block they were about to perform and also the reward values associated with each side of the display. Across the testing phase one side of the display (left or right hand targets) was consistently associated with a higher reward than the other. In the high reward blocks, the reward associated with the high rewarded side was £10, and the lower reward side was £1. In

the low reward blocks, the value for the highly rewarded side was £1, and the lower target side was £0.10. Left and right targets were equally likely to be presented.

*Stimuli*

The targets, letter "T" and the fixation square were the same as in Experiment 1 and outlined in the general method. In the testing phase there were two block types – low reward blocks and high reward blocks. The fixation square contained a single pound sign (£) in the low reward magnitude blocks and three pound signs (£££) in the high magnitude blocks. All pound signs were presented in red (15.8 cd/$m^2$).

These reward values (including the pound sign) were presented inside the corresponding target after a successful manual response to the target in each trial. The rewards were presented in a golden yellow colour (19.8 cd/$m^2$). If the participant was too slow, or made an incorrect response, no reward value was shown and a message appeared in the centre of the screen reading "Wrong! No reward" or "Too slow! No reward" (2.5$^o$ above fixation and white). The reward value (or error message) stayed on the screen for 1.5 seconds.

*Participants*

Eighteen participants (5 male) were recruited from the student population of the University of Bristol (approximate age range 18-25). All had normal or corrected-to-normal vision. Participants were reimbursed £7 for their time and received a variable performance related reward (£0, £0.1, £1 or £10) as outlined above.

*Design*

There were two within participant repeated measure binary factors: block (high or low reward) and side (high or low reward) leading to four conditions: high block, high side (£10 reward); high block, low side (£1); low block, high side (£1); low block, low side (£0.1).

The order of the blocks and the side with the high reward were counterbalanced using a latin square design across participants. The dependent measures were the manual response time and saccade latency.

*Manual Response Analysis*

A total of 5184 trials were recorded (288 x 18 participants). One participant was removed before further analysis because over 50% of their saccadic responses did not meet the fixation criteria applied. Given the magnitude of this percentage (68.4%), we decided to remove this participant from the manual analysis as well. We analysed the remaining manual reaction times excluding 839 time-out or response errors, leaving a total of 4131 data points. The distribution of errors is shown in Table 2. The range of the total percentage of time-out/response errors of each participant's responses was 3.1% - 36.1%.

*Saccadic Response Analysis*

We applied the same general exclusion criteria to initial fixation starting point as in Experiment 1. Trials were excluded if: the initial fixation was greater than 3 degrees from centre (11.9%, range across participants 0-33%); the saccades were in the incorrect direction (2.41%, range across participants 0-7.29%) or anticipatory (5.64%, range across participants 0-22.9%). The total number of remaining saccades was 3918 (80.0%). The remaining trials had a mean first saccade amplitude of 6.19 degree (range across participants 5.43 – 6.74), showing that all participants were, in general, making a large hyper-metric orienting saccade for their first saccade.

## Results

### *Manual Responses*

*Table 2.*

*Figure 3.*

The results are presented in Figure 3. Reward associated with the target side had a significant effect on manual reaction times leading to a 19.1ms (95% CI: 6.44 – 31.73) decrease (high condition) and 6.2ms (95% CI: -2.5 – 14.9) decrease (low condition) in response times for the high reward side (Repeated Measures ANOVA: [$F(1,16) = 7.8$, $p = 0.01$]; Effect size (partial ETA squared = 0.33). The effect of reward condition was not significant (Repeated Measures ANOVA: [$F(1,16) = 0.93$, $p = 0.35$]), but the interaction between the reward condition and target side was significant [$F(1,16) = 7.2$, $p = 0.02$; Effect size (partial ETA squared = 0.31)]. There was no significant effect of target side/reward condition on the frequency of errors (Repeated Measures ANOVA: Target Side: [$F(1, 16) = 2.59$, $p = 0.13$)], Reward Condition: [$F(1,16) = 0.49$, $p = 0.49$)]), or between the different types of errors (Table 2) [$F(1, 16) = 1.78$, $p = 0.2$].

Reaction times to the £1 target in the high reward condition (low side) are on average slower by 9.47ms (95% CI: -2.27 – 21.22ms) than to the target associated with the same reward in the low reward condition (high side) (see Figure 3), however this difference was not significant (Post-hoc Paired Samples t-test: [$t(16) = 1.7$, $p = 0.11$). This trend suggests that the effect of the reward values on reaction times may dependent on the context within which they are presented.

*Saccadic Responses*

There was no significant effect of reward condition (block) (Repeated Measures ANOVA: [F(1, 16) = 2.75, p = 0.12]) or reward side [F(1, 16) = 2.46, p = 0.14] but there was a significant interaction ([F(1, 16) = 13.45, p = 0.002]; Effect size (Partial ETA) = 0.46). Post-hoc testing on the high reward block in isolation shows a significant effect of reward side (Mean 125ms vs. 132ms, Paired Samples t-test: [t(16) = -2.769, p = 0.014]). This could indicate that although reward affects saccadic reaction times relatively (high vs. low within a block), the overall saliency of a highly rewarded condition has some influence. Further analysis showed no effect of reward condition or reward side on the frequency of anticipatory saccades; anticipatory in the incorrect direction (Repeated Measures ANOVA: Reward condition: [F(1, 16) = 0.7, p = 0.41], Reward side: [F(1, 16) = 1.31, p = 0.27]), anticipatory in the correct direction (Repeated Measures ANOVA: Reward condition: [F(1, 16) = 0.23, p = 0.64], Reward side: [F(1, 16) = 3.42, p = 0.08]).

To investigate if there was a relationship between individual performances across the two response modalities (saccadic and manual) we calculated the size of the reward effect for each of the 17 participants in the high and the low reward blocks separately. There was no strong correlation between the two response modalities: the Pearson's correlation r was 0.457 (p = 0.065) and 0.365 (p = 0.149) for the high and low reward conditions respectively.

## Experiment 2: Discussion

The effect of reward on both the manual and saccadic component of the response is consistent with the idea that reward processes mediated by dopamine neurons affect all

response systems irrespective of response modality (Schultz, 2010). However the clearest test of this association would be a strong correlation between the size of the reward effects across modalities. In the current experiment we found no such strong correlations. In other words, participants who showed a particularly strong reward effect in their saccadic responses were not necessarily the ones who showed a strong reward effect in their manual responses, or vice versa. This reflects the results we found in Experiment 1 for the probability manipulation.

In the low reward blocks in Experiment 2 there was a reduction in the difference between high and low reward targets for manual responses, and no evidence of an effect of reward for saccadic responses. However, there is a trend for faster reaction times to the £1 target when it is the high reward target (in the low reward blocks). This supports the evidence suggesting that the relative value of a stimulus that affects choice and saccadic reaction times (Milstein & Dorris, 2007, 2011) but also suggests there is a global reward component. The values of £1 and £0.10 are really quite low despite the relative difference between them, therefore global motivation could be lessened and this is likely to be why we see a smaller effect across modalities.

As discussed in the introduction, there is evidence to suggest a common system for reward processing that affects all motor systems to optimise the chance of getting the reward. Reward-related activity in the dopamine system initiates a series of events—one stage of which involves the anterior cingulate that eventually leads to changes in sensory representation (Hickey, Chelazzi, & Theeuwes, 2010a). Anterior cingulate and surrounding cortex is also known to be fundamentally involved in the control of attention and processing of attended stimuli (Hickey et al., 2010a; Hopfinger, Buonocore, & Mangun, 2000; Mesulam, 1999) A study using TMS (transcranial magnetic stimulation) found that the ACC (anterior cingulate cortex) facilitates implementation of a selected action, and is activated across three different output modalities (verbal, manual, oculomotor) (Paus, 2001). However, the

difference in the results reported here between the effect of reward on response times in manual and saccadic modalities might suggest otherwise. These differences could be explained by the effect that reward has on motor systems over time, if dopaminergic activation drives the reaction time effect. The dopamine reward signal is rapid and differs from the slower dopamine responses that have been associated with uncertainty, punishment and movement (Schultz, 2007). Activations in dopamine neurons to primary rewards, novel stimuli and reward-predicting stimuli have latencies of 60-100ms and endure for less than 200ms (Schultz, 2007). Given our results show very fast saccadic reaction times, it could be that the saccadic response is initiated too early to be strongly affected by the dopaminergic activation. This could be plausible if we assume that the dopamine activation occurs after the onset of the target given that the reward values are known. The manual responses however are much slower, and therefore could be affected more by the dopamine surge.

It is very interesting that this behavioural result on saccadic and manual responses is similar to what we see in the probability manipulation in Experiment 1. This supports the research discussed in the introduction of neurophysiological studies showing that dopamine neurons encode a combination of magnitude and probability (e.g. expected value) rather than distinguishing between the two (Tobler, Fiorillo & Schultz, 2005), and could suggest that the probability effect is more closely linked to reward processing even in the absence of primary rewards.

## Experiment 3: Combined reward and probability

To investigate the interactions between the effect of reward and probability on saccadic and manual responses, the next step was to combine these two factors within our paradigm to see if they combine in a simple manner. In Experiment 3, we investigate this across two response modalities to further understand the differences between these processes. As in Milstein and Dorris' (2007, 2011) study we manipulate the probability and reward associated with potential responses by varying the probability of the left vs. target presentation as well as the magnitude of reward associated with each. This allows us to vary the relative expected value of response to each target. This also reflects the paradigm in Experiment 1 (without the reward component) to allow for clearer comparison and continuity. In Experiment 3a, the reward values were kept constant and the probabilities manipulated across blocks, whereas in Experiment 3b the probability values were constant and the reward manipulated.

### Methods

#### *Procedure*

The task was the same as in Experiment 2. The experiment consisted of a practice phase (10 trials), a pre-test phase (48 trials), and then 9 blocks (48 trials each) of the testing phase. The reward schedule was random-lottery controlled, as in Experiment 2.

#### *Experiment 3a. Probability Manipulation*

The testing phase consisted of nine probability blocks. The participants were exposed to constant reward values in the testing phase, while the probability of targets being presented on either side was manipulated. Half of the participants were shown higher reward (£10) on the right and low reward on the left (£2) and the other half were shown the opposite. The

probabilities and reward values were changed from Experiments 1 and 2 to match EV. There were three experimental conditions, each consisting of three consecutive blocks. In one condition, the target was equally likely to be on the right or the left-hand side. In the second condition, the target was presented on one side with probability 0.83 and the other side with probability 0.17. These probabilities were then flipped spatially for the final condition. These probabilities and reward combinations lead to matched EV of £1.70 for two conditions: 0.83/£2 and 0.17/£10.

*Experiment 3b. Reward Manipulation*

The other eighteen participants were exposed to constant probabilities of the presentation of targets across the experiment, while the reward values associated with targets was manipulated. Nine participants were presented with targets appearing on the right-hand side with probability 0.83 and the left-hand side with probability 0.17, and the other nine were shown the opposite. Again there were three experimental conditions, each consisting of three consecutive blocks. In one condition, the target was associated with equal reward (£6) on the right and the left-hand side. In the second condition, the target presented on one side was associated with £10 reward and other side with £2. The reward values were then flipped for the final condition. As before, these probabilities and reward combinations lead to matched EV of £1.70 for two conditions: 0.83/£2 and 0.17/£10.

*Stimuli*

All stimuli apart from the fixation point in the testing phase were identical to Experiment 2, and the task required at the target was the same. The fixation square in the testing phase contained three pound signs (£££) as a fixation point, presented in red (15.8 $cd/m^2$).

*Participants*

Thirty-six participants (9 male) were recruited from the student population of the University of Bristol (approximate age range 18-25); eighteen assigned to Experiment 3a and likewise to 3b. All had normal or corrected-to-normal vision. Participants were reimbursed £7 for their time and received a variable performance related reward (£2, £6, or £10) as outlined above.

*Design*

There were two within-participant repeated-measure binary factors in both forms of the experiment: In the reward manipulation, one was block (£2/£10, £10/£2, £6/£6) and the other was side (0.83% probability and 0.17% probability). This led to 6 conditions: low probability, low reward; low probability, high reward; high probability, high reward; high probability, low reward; high probability, equal reward; low probability, equal reward. In the probability manipulation, one factor was block (83%/17%, 17%/83%, 50%/50%) and the other was side (£2 and £10). This led to 6 conditions: high reward, high probability; high reward, low probability; low reward, high probability; low reward, low probability; high reward, equal probability; low reward, equal probability. Therefore there were only 2 conditions in each form of the experiment that differed from the other: each was an isolated manipulation of reward or probability.

The orders of conditions were counterbalanced using a latin square design. The dependent measures were the manual response time and saccade latency.

*Manual Response Analysis*

*Experiment 3a*

There were a total of 7776 trials recorded in the probability manipulation experiment (432 x 18 participants) and of these 6872 manual responses were correct/before the time-out. The distribution of the remaining 904 time-out/response errors is shown in the table below. The range of the total percentage of time-out/response errors across all participants was 3.7% to 18.5%.

*Experiment 3b*

As in experiment 3a, the total number of trials recorded was 7776. Of these trials 719 manual responses were discarded due to an incorrect response or to the participant not responding in time. This left a total of 7057 trials. The distribution of errors across conditions is given in the table below. The range across participants of the total percentage of response/time-out errors was 2.5% to 16.7%.

*Saccadic Response Analysis*

*Experiment 3a*

The same exclusion criteria as the previous experiments were applied. Trials were excluded: where the initial fixation was greater than 3 degrees either side of the centre of the fixation box (16.3%; range across participants 1.39 - 39.1%); Saccade in the incorrect direction (1.97%; 0.46 - 3.94%) and anticipatory saccades in the correct direction (3.91%; 0.23 - 10.42%). An additional criterion was applied so that all trials where the first saccade was initiated after the participant's 'time-out' were discarded. Two participants had over 40% of trials excluded when these criteria were applied, so they were removed from the rest of the saccadic analysis. The remaining trials had a mean first saccade amplitude of 5.6 degrees (range across participants: 5.0 – 6.3 degrees). After removing the invalid, anticipatory and error saccades, the total number of analysed saccades was 5379 (77.82%).

*Experiment 3b*

The same exclusion criteria as the previous experiments was applied to initial fixation starting point: trials where the initial fixation was greater than 3 degrees either side of the centre of the fixation box (14.57%; range across participants: 1.85 – 40.97); saccades in the incorrect direction (3.47%; 0.23 – 9.95) and anticipatory saccades (6.36%; 1.16 - 10.88). All trials where the first saccade was initiated after the participant's 'time-out' were also excluded. Five participants had over 40% of trials excluded when these criteria were applied, so they were removed from the rest of the saccadic analysis. The remaining trials had a mean first saccade amplitude of 5.29 degrees (range across participants: 4.37 – 5.79 degrees). After removing the invalid, anticipatory and error saccades, the total number of analysed saccades was 4246 (75.61%).

## Results

### *Experiment 3a: Probability manipulation*

#### *Manual Responses*

*Table 3.*

There was a marginally significant effect of error type, with generally more time-out errors than response errors being present across participants (Repeated Measures ANOVA: [$F(1, 17) = 4.28$, $p = 0.05$]; Effect size (partial ETA) = 0.2).

*Figure 4.*

A significant effect of target probability was seen on the manual response times (Repeated measures ANOVA: [$F(2,17) = 26.6$, $p < 0.001$]; Effect size (partial ETA) = 0.61). This significant effect is not present between the equal and high probabilities, as shown by the large and overlapping confidence intervals (CI 95%: High = 444 – 489; Equal = 462 – 500), compared to a clear difference between these probabilities and the low probability (CI 95%: 484 – 529). The two conditions with matched expected values (low probability & high reward, high probability & low reward) produced significantly different reaction times, with a mean difference of 39.1ms (CI 95%: 24 – 54.2).

Although not significant, the data suggests a small effect of reward when the probability is equal across targets; this is clearly extinguished once the probabilities are unequal (Figure 4).

*Saccadic Responses*

The target probability had a significant effect on saccadic response times (Repeated Measures ANOVA: [$F(2, 15) = 25.39$, $p < 0.001$]; Effect size (partial ETA) = 0.63). Although the data shows a trend towards participants making faster saccades when the target is associated with high reward relative to low reward (Figure 4), this effect is not significant [$F(1, 15) = 1.76$, $p = 0.21$]. The effect of probability is only present between the low probability condition (CI 95%: 138 – 162.5) and the high/equal probability condition (CI 95%: High = 127 – 143; Equal = 128 – 144), as can be clearly seen in the graph (Figure 4). When the Expected Value were matched across the two targets, (the ‘high reward, low probability’ and the ‘low reward, high probability’ points in Figure 4) there was a 9.63ms (CI 95%: -1.48 – 20.73) increase in response time between the high probability side and the low

probability side. This result suggests that the processes underlying the reward and probability effect are not governed by EV.

There was a significant interaction between the percentages of correct anticipatory saccades with reward side and probability condition across participants (Repeated measures ANOVA: [$F(2, 15) = 8.69$, $p = 0.001$]; Effect size (Partial ETA) = 0.37). There was no significant effect of reward side or probability condition on the percentage of correct anticipatory saccades (Reward side: [$F(1, 15) = 1.23$, $p = 0.29$], Probability Condition: [$F(2, 15) = 0.92$, $p = 0.41$]). Additionally, there was no effect of either reward or probability on the percentage of anticipatory saccades in the incorrect direction (Reward side: [$F(1, 15) = 0.03$, $p = 0.9$], Probability Condition: [$F(2, 15) = 0.3$, $p = 0.74$]) but there was a significant interaction ([$F(2, 15) = 11.8$, $p < 0.001$]; Effect size (Partial ETA) = 0.44).

*Experiment 3b: Reward manipulation*

*Manual Responses*

*Table 4.*

There was a significant effect of probability on both the frequency of time-out and response errors made, where more errors were made overall in the low probability direction (Time-out errors: Repeated measures ANOVA: [$F(1, 17) = 36.16$, $p < 0.001$]; Effect size (Partial ETA) = 0.68; Response errors: Repeated measures ANOVA: [$F(1, 17) = 7.8$, $p = 0.012$]; Effect size (Partial ETA) = 0.32).

*Figure 5.*

Target probability had a significant effect on manual responses: response times were shorter for the higher probability side (Repeated Measures ANOVA: [F(1, 17) = 39, p <0.001]; Effect size (partial ETA) = 0.7). The reward manipulation had no significant effect on the manual responses across participants [F(2, 17) = 2.33, p = 0.11)]. There was a significant difference between the conditions for which the expected value was matched across the two targets, with a mean difference of 35.26ms (CI 95%: 15.14 – 55.37); this suggests EV did not govern manual responses.

*Saccadic responses*

There was no significant effect of target probability or reward on saccadic response times (Repeated Measures ANOVA; Reward condition: [F(2, 12) = 0.59, p = 0.56], Probability: [F(1, 12) = 3.53, p = 0.09]). The data showed a trend towards an effect of probability in the high and equal reward conditions, although this was not evident in the low reward condition (Figure 4).

The correct anticipatory saccade trials were analysed by frequency within reward condition and probability sides, and the percentages of the total possible trials were calculated for each participant. There was a significant effect of probability on the percentage of correct anticipatory saccades, whereby more anticipatory saccades were made towards the higher probability target (Repeated Measures ANOVA: [F(1, 12) = 14.33, p = 0.003]; Effect size (partial ETA) = 0.54). There was no effect of reward on the percentage of correct anticipatory saccades [F(2, 12) = 1.01, p = 0.38], nor an effect of either reward or probability on the percentage of anticipatory saccades in the incorrect direction (Reward: [F(2, 12) = 0.8, p =

0.46], Probability: [F(1, 12) = 2.75, p = 0.12]). As with the manual results there was no evidence of expected value governing saccadic response times, as the reaction times across matched expected value conditions were significantly different (Mean = 10.97, CI 95%: 0.49 – 21.45).

## Discussion

Across Experiments 3a and 3b there was a consistent effect of target probability on manual and saccadic responses: participants' response times were faster to targets associated with higher probability, as found in Experiment 1. However, there was no significant effect of probability on the saccadic responses in Experiment 3b. There was a trend for faster responses to higher probability and slower responses to lower probability in the high and equal reward conditions, but this effect was completely extinguished in the low reward condition. In Experiment 3a, the probability effect was only present between the high/equal and the low probability conditions.

There was no significant effect of reward in either Experiment 3a or 3b, and no evidence to suggest that EV could predict either saccadic or manual responses. In the manual results of Experiment 3a, some evidence was found to suggest that a small reward effect was present only when probability was equal across targets.

These results suggest that the reward effect is significantly modulated when there are unequal probabilities of the targets that rewards are associated with. It could be that the strength of the probability effect completely dominates the modulation in behaviour and thus these processes do not combine in a simple manner. Our results suggest that there are differences between the two processes governing reward and probability, but that they do not act completely independently on behaviour.

## Experiment 4: Isolated reward and probability

We next report the final experiment investigating the effects of reward and probability in isolation, using the same paradigm for both effects and a within-subjects design. This was to allow for a direct comparison of the reward and probability effect without combining them together, and by running both isolated conditions on the same participants, to examine whether the two effects are correlated. Given the smaller and less reliable effect of probability and reward on saccadic responses in comparison to manual responses in the last three experiments, only manual responses have been recorded in Experiment 4. Additionally, the same two keys were used to respond to targets on the left and right-hand side of the screen so that participants only used one hand to respond. In the previous experiments it is not clear at what stage in processing the manual response advantages reported previously occur. As each side was unequally associated with a particular hand it is entirely possible that the speeding up was a result of some low-level motor readiness associated with one hand rather than the other. By having both responses (left or right) being made by the same hand this allows a test of whether these effects are more central than this explanation suggests.

### Methods

#### *Procedure*

The task layout was the same as the last three experiments, except that regardless of which hemifield the target was presented in, the same two buttons on the keyboard were used to indicate the orientation of the letter T. The experiment consisted of a practice phase (10 trials), a pre-test phase (36 trials), and then 12 blocks (36 trials each) of the testing phase. The same reward schedule was employed as in Experiment 3.

*Stimuli*

All stimuli and the fixation square were the same as the previous experiments. The fixation square contained a single cross presented in red (15.8 cd/$m^2$) and in font size 17. The testing phase was split into two sections – the reward manipulation and the probability manipulation. In the reward manipulation the target was presented on either side in an equal number of trials across all conditions, whereas the frequency changed in different conditions of the probability manipulation. Half the participants were exposed to the reward conditions first, and half to the probability. In the reward manipulation, there were three conditions: high reward right (£2/£10), high reward left (£10/£2), and equal reward (£6/£6). In the probability manipulation, there were again three conditions: high probability right (0.17/0.83), high probability left (0.83/0.17) and equal probability (0.5/0.5). Each condition consisted of two consecutive blocks of 36 trials. These probabilities were chosen as they are the closest match to £10/£2 reward proportions. To ensure motivation within the probability conditions was not decreased compared to the reward conditions, rewards were fixed at £6 across the whole section. This meant that the equal reward and equal probability conditions were exactly the same. Feedback on incorrect/slow trials was given as in Experiment 3.

*Participants*

Eighteen participants (eleven female) were recruited from the student population of the University of Bristol (mean: 22.8, range: 18 - 40). All had normal or corrected-to-normal vision. Participants were reimbursed £6 for their time and received a variable performance related reward (£0, £2, £6 or £10) as outlined above. They were asked to use their self-reported dominant hand to do the experiment (14 right-handed). The study was approved by The Faculty of Science Human Research Ethics Committee at the University of Bristol.

*Design*

There were two within participant repeated measure factors: probability (high, low or equal) and reward (high, low or equal) leading to six conditions: High probability, low probability, equal probability; High reward, low reward, equal reward. The orders of the blocks were counterbalanced using a latin square design across participants, and participants were told which condition they were in at the start of each block. The dependent measure was the manual response time.

*Manual Response Analysis*

A total of 7776 trials were recorded (432 x 18 participants). For the manual reaction times we excluded 588 time-out and 470 response errors, leaving a total of 6718 data points. The distribution of errors is shown in Table 5. The range of the total percentage of time-out/response errors of each participant's responses was 0.9% - 27.78%.

Results

*Manual Responses*

*Table 5.*

*Figure 6.*

We found that reward value had a significant effect on manual reaction times, and exhibited a linear relationship across the three conditions (Figure 6) (Repeated Measures ANOVA: [$F(2, 17) = 7.55$, $p = 0.002$]; Effect size (partial ETA squared) = 0.31). Probability also had a significant effect on manual reaction times, and a linear relationship can also be seen in Figure 6 (Repeated Measures ANOVA: [$F(2, 17) = 28.03$, $p < 0.001$]; Effect size (partial ETA squared) = 0.62). There was no difference between reward and probability in the equal condition (Figure 6): these conditions were identical besides from the participants experiencing them within the reward or probability manipulation of the experiment.

There were significantly more time-out errors made the lower the probability or reward value was across the experiment (Probability: Repeated Measures ANOVA: [$F(2,17) = 27.03$, $p < 0.001$]; Effect size (partial ETA squared) = 0.61; Reward: Repeated Measures ANOVA: [$F(2,17) = 6.13$, $p = 0.005$]; Effect size (partial ETA squared) = 0.27). However, there was no effect of reward or probability on the frequency of response errors (Probability: [$F(2, 17) = 0.84$, $p = 0.44$], Reward: [$F(2, 17) = 2.79$, $p = 0.08$])

There was only a weak non-significant positive correlation between the effect of reward and probability across all participants (Figure 6, right hand panel; correlation coefficient = 0.24, N.S. [p=0.34]). Because the response was now uni-manual, and so lateralized, we next investigated if these was an effect on the side of presentation of the target. The data was split for the 14 right-handed participants into manual responses made to the left or the right-hand target. We ran two repeated measures ANOVAs for both reward/probability conditions, and found that there was no significant difference between manual responses made to the left or

right side target (Repeated Measures ANOVAs: Left/right side of probability condition: [$F(1, 17) = 0.01$, $p = 0.92$], or for the left/right side of reward condition: [$F(1, 17) = 3.12$, $p = 0.1$]).

## Experiment 4: Discussion

Reward and probability had a significant linear effect on the manual reaction times, whereby responses were faster to targets associated with higher reward and higher probability. This modulation in responses was larger in the probability manipulation. It was found that the lower the reward or probability, the greater the frequency of time-out errors; this suggests, as discussed in the introduction, that both high reward and probability could lead to increased motor preparation (Basso & Wurtz, 1997; Dorris & Munoz, 1998; Koval et al., 2004). Additionally, it was found that there was no difference between responses made to targets on the right or the left-hand side of the screen. It could be concluded from this that spatial mapping of visual stimuli to manual responses is not an important aspect of the effect of reward and probability on reaction times.

The results of Experiment 4 show that the effect of reward on manual responses is smaller than the effect of probability. One factor that could affect this is the likelihood of individual differences in sensitivity to reward. A number of studies have identified this, for example Hickey, Chelazzi, and Theeuwes (2010b) measured trait reward-seeking using a personality index, and found this correlated with the magnitude of reward priming in a visual search task. The term “reward priming” was used to describe the bias towards selection of objects previously characterised as rewarding (Hickey et al., 2010b). They also found in a similar study that the ERP component known to be a sufficient index for reward processing in the anterior cingulate cortex is elicited during reward feedback processing, and that the magnitude of this predicts the effect of reward on each participant’s behaviour during visual

search (Hickey et al., 2010a). Linking primary explicit reward to a personality trait and linking it to a single neural system would suggest that the magnitude of the reward effects across response types should be correlated across participants. If a participant is particularly reward sensitive this sensitivity should be expressed both in their saccadic and their manual responses. In Experiments 2 and 3 there is no evidence in favour of such a correlation, despite having quite large variability in reward sensitivity across participants and finding reliable reward effects overall. However, we should be wary of making conclusions about individual differences with a relatively small sample size.

There was no correlation between reward and probability in our data, this also supports the results of Experiment 3 where we did not see these processes combining in a simple manner. This could be partly due to the differences in the nature of reward and probability; the probability of events or objects appearance in our environment is inherent and continues to exist (albeit fluctuating) even if not observed. The explicit reward value attributed to an event or object relies on the way in which we perceive it and is thus dependent on many other factors that can differ between individuals (utility, internal state etc). Several authors have suggested information acquired at fixation can be thought of as a secondary (implicit) reward and can mediate gaze-learning patterns (Tatler, Hayhoe, Land, & Ballard, 2011; Jovancevic-Misic & Hayhoe, 2009). Paeye, Schütz and Gegenfurtner (2016) have showed that finding a target (therefore acquiring information) at a location can reinforce eye movements in visual search. Therefore, if probability dominates when combined with explicit reward it may be that there is an intrinsic reward component from exploration of the target that is stronger than explicit reward in isolation. This would mean that higher probability is associated with higher implicit reward. This is supported by the results of Experiment 4 which show the isolated probability effect to be greater and more robust within participants than explicit reward. One way to formally investigate this is to look at whether the simple combining of reward and

probability as Expected Value (probability of rewarded event X magnitude of the reward, as discussed in the introduction) can explain the data across our four experiments.

## Modelling the Data

The Expected Value (EV) at a particular location (n) is given by:

$$EV_n = P_n R_n \quad (1)$$

Where *P* is the probability of a rewarded event in that location and *R* is the reward at that location.

In the experiments in this paper there are only two locations which we will label *1* and *2* so Relative Expected Value (REV) at location 1 is given by:

$$REV_1 = \frac{EV_1}{EV_1 + EV_2} \quad (2)$$

Note also that:

$$REV_1 = 1 - REV_2 \quad (3)$$

$$P_1 = 1 - P_2 \quad (4)$$

We want to determine the extent to which the *REV* determines the response times (*RT*) observed across experiments so we also calculate a simple metric of the relative response time differences in a given condition where there are two possible response locations. We chose to use the percentage difference (*PD*) in reaction time which we define as:

$$PD = 100 \frac{RT_1 - RT_2}{\overline{RT}} \quad (5)$$

As $REV_1$ is dependent on $REV_2$ (see equation 3) we only plot the data from the location with the higher *REV*.

For Experiment 1 the Reward was 0 so REV is undefined across these conditions. For the remaining conditions we can plot and correlate REV and PD. Figure 7 plots this relationship. The correlation coefficient is low (r=0.174). This simple formulation of REV does not appear to capture the variability in the response time data well. An explanation for this might be due to the difference in our experimental design compared to Milstein and Dorris' (2007) study. They employed a paradigm whereby reward is associated with every single trial and built up in small increments. This means for example than an incorrect/slow response to a high probability but low reward trial means overall reward is reduced by a relatively small amount, rather than increasing the likelihood of receiving £0 reward overall (as in our series of experiments). Essentially the stakes are higher per trial in our study. EV does not explain our data well.

Having collected the data for the series of experiment reported here, and as a result of the poor fits that we obtained for REV we became aware of an argument in the literature for there being an intrinsic reward associated with exploration and information acquisition (see Collins, 2012; Deaner, Khera, and Platt 2005; Tatler et al., 2011) as discussed. As a result we carried out a further post-hoc modelling exercise, which is an extension of the REV model.

Within this extension of the REV model the total reward associated with a location consists of the explicit reward associated with that location (*E*) plus an intrinsic reward associated with exploration of that location (*I*).

$$R_n = E_n + I_n \quad (6)$$

For our experiments, *I* will be a function of the probability of the rewarded target occurring at that location (*P*) multiplied by some weighting factor (*k*) to scale the effects of the explicit and intrinsic rewards.

$$R_n = E_n + (kP_n) \quad (7)$$

Combining (1) and (7) gives:

$$EV_n = P_n(E_n + kP_n) \quad (8)$$

Combining (2) and (8) gives:

$$REV_1 = \frac{P_1(E_1+kP_1)}{P_1(E_1+kP_1) + P_2(E_2+kP_2)} \quad (9)$$

We label this new measure of Relative Expected Value as REV+. We next carried out a linear fit between PD and REV+ for each condition across all the experiments fitting a single value of *k* as a free parameter. In this fitting procedure, we exclude Experiment 1 as before.

*Figure 7.*

The fit is extremely good with a correlation coefficient r=0.95. These data are plotted as disks in Figure 7 (right hand graph). The fit estimates k as 17.31. In other words if the target is equally likely to come up on both sides (i.e. *P*=0.5) and there is no Explicit Reward (*E*) then the Implicit Reward is equivalent to £9.17 which is close to the highest Explicit Reward value that we used across the experiments (£10). We can now use the model to generate predictions for the results from Experiment 1 given that there was no explicit reward. These results are shown as crosses in Figure 7 (right hand graph) and appear to be an excellent fit to the model as well.

Implicit reward clearly plays a large part in determining behaviour in these experiments. However, note that explicit reward still has an important part to play. Explicit reward still modulates responses across Experiments 2 and 4 when the probability (and

therefore implicit reward) is the same across the targets. These data are well fitted by the model highlighting the contribution explicit reward plays in the model fits.

Although we see weaker effects of reward and probability on saccadic responses across the first three experiments, our model is still a good fit to the saccadic response data ($r$ = 0.84) particularly when compared to the simple Expected Value model ($r$ = -0.01).

## General Discussion

### *Response Modalities*

We replicated the probability effect established in the literature across the four experiments, however this effect was weaker in the saccadic responses and we find no evidence that it was correlated across response modalities. The results of the experiments (2-4) on reward manipulations showed a stronger effect of reward on manual than saccadic responses. Both the effects of reward and of probability on different response modalities seem to be sensitive to the time course of the stimulus presentation. In the literature, stronger probability effects are seen on saccadic responses when the ISI is shorter. Additionally, given our model which supports an implicit reward component related to probability, in general the stronger effect on manual responses across Experiments 1 to 3 could be linked to the time course of dopamine activation (Schultz, 2007). Activations in dopamine neurons to reward-predicting stimuli are very fast and have latencies of 60-100ms and endure for less than 200ms. Fast saccadic responses may be initiated too early to be affected as strongly by the dopaminergic activation, when compared to longer latency manual responses.

Saccadic reaction times may also have been reduced by the presence of a perceptual task at the saccadic target and could explain why there was no correlation between response

modalities. The inclusion of a perceptual task at the landing point of the saccade clearly constitutes a form of information worth gathering as a correct response leads to receiving reward. Supporting this Montagnini and Chelazzi (2005) found that a perceptual task at a saccadic target reduces reaction times by up to 15% and more recently Wolf and Schütz (2017) found saccadic reaction times were reduced by up to 40ms by the performance of a perceptual task at the target (when task relevant visual information is foveated). The fast saccadic responses seen across the four experiments could constitute a 'floor effect' whereby little room is left for further modulations by reward and probability.

Perhaps more pertinent to the differences across response modalities is the effect of negative feedback on manual response speed and accuracy. Across all experiments, a stronger effect of probability/reward was seen on manual responses than saccadic. The motivation to respond is more clearly focussed on the manual response to the target, which is explicitly goal-oriented given that the button press required is the primary task for the participants. This is then directly linked to the speed and accuracy of the response and thus the feedback received on each trial. This could explain why we see a stronger effect of reward and probability on manual responses compared to saccadic: the purpose of the saccade is primarily to gather information (Hickey et al., 2010a). With regards to this point, and the absence of correlation between saccadic and manual responses across all experiments, it is unlikely that both response modalities could be explained by a decision-making model where all signals across pathways are brought together with different decision thresholds for different responses (Bompas & Sumner, 2008). However, it is important to note that the difference in the purpose and execution of the saccadic and manual response in these experiments could be the significant factor that sets the two response modalities apart and leads to the lack of evidence for a correlation. As discussed in the introduction, the pre-motor theory of attention suggests a common source of information is used for all types of motor

responses and that the same decision threshold would apply to all responses (Bompas & Sumner, 2008). The decision threshold for the saccadic latency in the experiments reported here relates to the side on which the target has appeared. Conversely, the decision threshold for the manual response additionally encapsulates the time taken to reach a binary perceptual decision at the location of the target; this is unrelated to the spatial position of the target. Thus it follows that these two responses may be too different to be able to combine as a common source of information. Further work investigating both response modalities in a paradigm where the perceptual task requires a similar goal-oriented fast and accurate saccadic and concurrent manual response (e.g. a saccade up or down to indicate the orientation of a letter T) could clarify this issue.

*Probability, Explicit and Intrinsic Reward*

Our simple single parameter model provides a very good fit to a wide range of data across the manual response data and the saccadic data. Even in the absence of explicit primary rewards (such as in Experiment 1) the predictions of the model fit the data well. Without the presence of primary rewards, the purpose of eye movements is to obtain information, which is used to achieve behavioural goals (Tatler et al., 2011). The behavioural goal in our series of experiments would be the manual response, which is then directly linked to the obtaining of information. This acquisition of information and subsequent reaching of behavioural goal can be thought of as a secondary reward (Tatler et al., 2011). This is supported by a model by Sprague and Ballard (2003), which assumes eye movements are valued by estimating the expected cost of the uncertainty that will result if the movement is not made. Therefore, reward is maximised by allocating implicit reward to the behaviour that stands to lose the most (Sprague & Ballard, 2003). If we consider the example of a high and

low probability manipulation, the higher probability location will always provide the most information and thus might be prioritised over the low probability location regardless of the explicit primary reward. This has been researched explicitly in the saccadic domain, in a study where participants learnt that targets at one location were more likely to remain visible after a successful saccade than targets at other locations (Collins, 2012). Saccades to these targets were faster, more frequent and harder to inhibit than saccades to other targets. This increase in saccadic readiness towards locations more likely to lead to increase in visual information suggest post-saccadic visual feedback acts an implicit reward. Collins (2012) suggests that sensory-motor systems have their own intrinsic reward, and that informational content is likely to be a more relevant "currency" for the oculomotor system than monetary or other explicit primary reward. This informational content may simply be a clearer view of visual objects, and this in itself may be rewarding and contribute to on-going movement monitoring in learning (Collins, 2012). These studies and the results of our experiments further support the theory that eye movement decision are being more affected by biases in how we move our eyes rather than those biases contributing to image salience (Tatler et al., 2011).

Research on non-human primates has quantified the intrinsic reward associated with looking at particular visual stimuli (Deaner, Khera, & Platt, 2005; Shepherd, Deaner, & Platt, 2006). One study found that rhesus macaques differentially value the opportunity to acquire visual information about members of their social group, modulated by their social status (Deaner, Khera, & Platt 2005). The monkeys sacrificed liquid reward for visual information about higher status individuals, and required overpayment of juice to view lower social status members (Deaner, Khera, & Platt 2005). Here they have used liquid as a measurable explicit equivalent of an implicit reward resulting from gaze. This supports our suggestion that looking in itself provides an implicit reward; dopaminergic machinery appears to be

intimately related to the sensitivity of eye movement target selection to behavioural outcomes. Further support for the influence of implicit reward on the dopaminergic system comes from a model of reward processing in the basal ganglia in humans, where the contrast between reinforcement probabilities is enhanced by subtracting "Go" learning (speeding up responses for higher probabilities) and "NoGo" learning (slowing down responses to lower probabilities) associations but large reward magnitudes are underweighted (Frank, 2005; Frank & Claus, 2006; cited in Moustafa, Cohen, Sherman, & Frank, 2008).

The lesser influence of explicit reward in our studies could also be explained by the paradigm used, specifically the absence of dual-target (choice) trials. Milstein and Dorris (2007, 2011) found that during dual-target trials, the higher reward magnitude target was more frequently chosen. Latencies in single-target trials were more affected by expected value than by reward magnitude alone; this contradicts our results. Given that these different trial types were interleaved within the same experiment, an interaction may have affected the results (Wolf, Heuer, Schubö, & Schütz, 2017). Wolf et al (2017) found that having dual target (choice) trials interleaved in blocks modulated saccadic and manual latencies to differential reward values in single target trials, but this modulation was mostly absent in blocks with no dual target trials. In light of our findings, it's possible that the weaker effects of reward seen are because reward is more likely to affect preparation of responses when it is more behaviourally relevant, as when choice trials are present. In these instances it is more beneficial to prioritise the more rewarding target/side.

The results of our four experiments indicate the importance of considering implicit reward in behavioural reward studies. Different types of bias that affect our response to stimuli in the environment rarely combine in a simple manner, particularly biases like explicit (primary) reward, that can vary hugely in 'currency' and are dependent on the goals of the individual. Our results suggest that reward may have an implicit component that comes from

gaining information and achieving task-related goals. The importance of secondary implicit rewards is particularly clear when we consider that primary rewards (even more so monetary) rarely happen in real life from a simple eye movement or hand movement. Acquiring information and achieving task goals are the most likely real-life purpose of the many saccades and hand movements made during this task, which then ultimately lead to primary rewards that are important for survival. Our findings suggest that a different approach to reward related studies in humans, rather than just providing primary monetary reward, could further our understanding of the governing processes and reward circuitry.

## Acknowledgements

We thank Chris Kent for very helpful comments on an earlier version of this manuscript and members of the Bristol Decision Making Group for ongoing discussion about the work. This work was financially supported by two grants from the UK Engineering and Physical Sciences Research Council (EP/I032622/1, EP/M000885/1).

## Declaration of Interest

The authors report no potential conflict of interest.

References

Antoniades, C. A., Bogacz, R., Kennard, C., FitzGerald, J. J., Aziz, T., & Green, A. L. (2014). Deep brain stimulation abolishes slowing of reactions to unlikely stimuli. *The Journal of Neuroscience*, *34*, 10844-10852

Basso, M. A., & Wurtz, R. H. (1997). Modulation of neuronal activity by target uncertainty. *Nature*, *389*, 66-69.

Bendiksby, M. S., & Platt, M. L. (2006). Neural correlates of reward and attention in macaque area LIP. *Neuropsychologia*, *44*, 2411-2420.

Bompas, A., & Sumner, P. (2008). Sensory sluggishness dissociates saccadic, manual, and perceptual responses: An S-cone study. *Journal of Vision*, *8*:10, 1-13

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433-436.

Collins, T. (2012). Probability of seeing increases saccadic readiness. *PloS one*, *7*, e49454.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, *1*, 115-131.

Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Current Biology*, *15*, 543-548.

Dorris, M. C., & Munoz, D. P. (1998). Saccadic probability influences motor preparation signals and time to saccadic initiation. *The Journal of Neuroscience, 18*, 7015-7026.

Dreher, J. C., & Tremblay, L. (Eds.). (2009). *Handbook of Reward and Decision Making*: Academic Press.

Druker, M., & Anderson, B. (2010). Spatial probability aids visual stimulus discrimination. *Frontiers in Human Neuroscience*, *4*, 63.

Dunne, S., Ellison, A., & Smith, D. T. (2015). Rewards modulate saccade latency but not exogenous spatial attention. *Frontiers in Psychology, 6*, 1080.

Eimer, M., Van Velzen, J., Gherri, E., & Press, C. (2006). Manual response preparation and saccade programming are linked to attention shifts: ERP evidence for covert attentional orienting and spatially specific modulations of visual processing. *Brain Research, 1105*, 7-19.

Fecteau, J. H. & Munoz, D. P. (2006). Salience, relevance, and spiking neurons: a priority map governs target selection. *Trends in Cognitive Sciences, 10*, 382-390.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience, 17*, 51-72.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological Review, 113*, 300.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010a). Reward changes salience in human vision via the anterior cingulate. *The Journal of Neuroscience, 30*, 11096-11103.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010b). Reward guides vision when it's your thing: Trait reward-seeking in reward-mediated visual priming. *PLoS One, 5*, e14087.

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience, 3*, 284-291.

Jóhannesson, Ó. I., Haraldsson, H. M., & Kristjánsson, Á. (2013). Modulation of antisaccade costs through manipulation of target-location probability: Only under decisional uncertainty. *Vision Research, 93*, 62-73.

Jones, J. L., & Kaschak, M. P. (2012). Global statistical learning in a visual search task. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 152.

Jovancevic-Misic, J., & Hayhoe, M. (2009). Adaptive gaze control in natural environments. *Journal of Neuroscience, 29*, 6234-6238.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*, 263-291.

Kalesnykas, R. P., & Hallett, P. E. (1994). Retinal eccentricity and the latency of eye saccades. *Vision Research, 34*, 517-531.

Kaufman, H., & Levy, R. M. (1966). A further test of Hick's law with unequally likely alternatives. *Perceptual and Motor Skills, 22*, 967-970.

Kawagoe, R., Takikawa, Y., & Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience, 1*, 411-416.

Körner, C., & Gilchrist, I. D. (2007). Finding a new target in an old display: Evidence for a memory recency effect in visual search. *Psychonomic Bulletin & Review, 14*, 846-851.

Koval, M. J., Ford, K. A., & Everling, S. (2004). Effect of stimulus probability on anti-saccade error rates. *Experimental Brain Research, 159*, 268-272.

Lamb, J., & Kaufman, H. (1965). Information transmission with unequally likely alternatives. *Perceptual and Motor Skills, 21*, 255-259.

Liston, D. B., & Stone, L. S. (2008). Effects of prior information and reward on oculomotor and perceptual choices. *The Journal of Neuroscience, 28*, 13866-13875.

Liu, C. L., Chiau, H. Y., Tseng, P., Hung, D. L., Tzeng, O. J., Muggleton, N. G., & Juan, C. H. (2010). Antisaccade cost is modulated by contextual experience of location probability. *Journal of Neurophysiology, 103*, 1438-1447.

Liu, C. L., Tseng, P., Chiau, H. Y., Liang, W. K., Hung, D. L., Tzeng, O. J. L.,…Juan, C. H. (2011). The location probability effects of saccade reaction times are modulated in the frontal eye fields but not in the supplementary eye field. *Cerebral Cortex, 21*, 1416-1425.

Mesulam, M. M. (1999). Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *354*, 1325-1346.

Milstein, D. M., & Dorris, M. C. (2007). The influence of expected value on saccadic preparation. *The Journal of Neuroscience, 27*, 4810-4818.

Milstein, D. M., & Dorris, M. C. (2011). The relationship between saccadic choice and reaction times with manipulations of target value. *Frontiers in Neuroscience, 5*, 122.

Montagnini, A., & Chelazzi, L. (2005). The urgency to look: prompt saccades to the benefit of perception. *Vision research, 45*(27), 3391-3401.

Moustafa, A. A., Cohen, M. X., Sherman, S. J., & Frank, M. J. (2008). A role for dopamine in temporal decision making and reward maximization in parkinsonism. *The Journal of Neuroscience*, *28*, 12294-12304.

Mowbray, G. H. (1964). Subjective expectancy and choice reaction times. *Quarterly Journal of Experimental Psychology*, *16*, 216-223.

Nakahara, H., Nakamura, K., & Hikosaka, O. (2006). Extended LATER model can account for trial-by-trial variability of both pre-and post-processes. *Neural Networks, 19*, 1027-1046.

Noorani, I., & Carpenter, R. H. S. (2013). Antisaccades as decisions: LATER model predicts latency distributions and error responses. *European Journal of Neuroscience, 37*, 330-338.

Paeye, C., Schütz, A. C., & Gegenfurtner, K. R. (2016). Visual reinforcement shapes eye movements in visual search. *Journal of vision*, *16*(10):15, 1-15.

Paus, T. S. (2001). Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nature Reviews Neuroscience*, *2*, 417-424.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. *Attention and Performance XV*, *15*, 231-265.

Schultz, W. (2007). Multiple dopamine functions at different time courses. *Annual Review of Neuroscience*, *30*, 259-288.

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behavioural and Brain Funct*ions, *6*, 1-9.

Shepherd, S. V., Deaner, R. O., & Platt, M. L. (2006). Social status gates social attention in monkeys. *Current Biology*, *16*, R119-R120.

Sprague, N., & Ballard, D. (2003). Multiple-goal reinforcement learning with modular sarsa (0). In *International Joint Conference on Artificial Intelligence* (pp.1445-1447). San Francisco, CA: Morgan Kaufmann.

Stankevich, B. A., & Geng, J. J. (2015). The modulation of reward priority by top-down knowledge. *Visual Cognition*, *23*, 206-228.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*, 403-409.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: reinterpreting salience. *Journal of Vision*, *11*, 1–23.

Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., & Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Experimental Brain Research, 142*, 284-291.

Taylor, M. J., Carpenter, R. H. S., & Anderson, A. J. (2006). A noisy transform predicts saccadic and manual reaction times to changes in contrast. *The Journal of Physiology, 573*, 741-751.

Theeuwes, J., & Belopolsky, A. V. (2012). Reward grabs the eye: Oculomotor capture by rewarding stimuli. *Vision Research, 74*, 80-85.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science, 307*, 1642-1645.

Walthew, C., & Gilchrist, I. D. (2006). Target location probability effects in visual search: an effect of sequential dependencies. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1294.

Wolf, C., Heuer, A., Schubö, A., & Schütz, A. C. (2017). The necessity to choose causes the effects of reward on saccade preparation. *Scientific reports, 7*(1), 16966.

Wolf, C., & Schütz, A. C. (2017). Earlier saccades to task-relevant targets irrespective of relative gain between peripheral and foveal information. *Journal of vision, 17*(6):21, 1-18.

## Tables

| Target probability | | | | | |
|---|---|---|---|---|---|
| 0.75 | | 0.5 | | 0.25 | |
| Time-Out | Response Errors | Time-Out | Response Errors | Time-Out | Response Errors |
| 5.52 (2.78-9.03) | 4.05 (0-9.72) | 6.02 (1.04-14.58) | 4.57 (0-16.67) | 8.33 (0-22.92) | 5.44 (0-16.67) |

*Table 1. the mean percentage of errors (of the total number of possible trials in each condition), for both types; time-outs and response errors. The range across 18 participants is given in parentheses.*

| Side | Block Type | | | |
|---|---|---|---|---|
| | High Block | | Low Block | |
| | Time-Out | Response Errors | Time-Out | Response Errors |
| High Reward | 6.62 (0-13.9) | 7.19 (0-20.8) | 7.92 (1.40-19.4) | 6.78 (0-20.8) |
| Low Reward | 9.89 (1.39-34.7) | 8.25 (1.39-19.4) | 9.72 (2.77-25.0) | 6.54 (0-19.4) |

*Table 2. The mean percentage of errors for Experiment 2, for both error types: time-outs and response errors. The range across 17 participants is given in parentheses*

| **Side** | **Block Type** | | | | | |
|---|---|---|---|---|---|---|
| | Equal Probability | | Probability same direction as reward | | Probability opposite direction to reward | |
| | Time-Out | Response Errors | Time-Out | Response Errors | Time-Out | Response Errors |
| High Reward | 7.1 (2.78 - 23.6) | 5.25 (0 - 9.7) | 4.81 (0.83 - 13.3) | 4.58 (0.83 - 7.5) | 6.94 (0 -29.17) | 6.94 (0 -16.67) |
| Low Reward | 8.1 (0 -20.83) | 4.5 (0 -22.2) | 10.42 (0 - 25) | 5.79 (0 - 16.67) | 5.93 (0 - 14.17 | 3.98 (0 - 9.17) |

*Table 3: Mean percentage of errors for Experiment 3a across the three probability conditions and two reward sides. Errors are split into time-out errors and response errors. The range of error percentages across participants in each condition are given in brackets.*

| **Side** | **Block Type** | | | | | |
|---|---|---|---|---|---|---|
| | Equal Reward | | Reward same direction as probability | | Reward opposite direction to probability | |
| | Time-Out | Response Errors | Time-Out | Response Errors | Time-Out | Response Errors |
| High Probability | 3.7 (0 – 8.3) | 4.07 (0.83 - 10) | 3.66 (0 – 9.17) | 3.8 (0 - 10) | 3.98 (0 – 9.17) | 4.07 (0 – 12.5) |
| Low Probability | 9.95 (0 – 16.7) | 6.25 (0 - 25) | 10.42 (0 – 29.17) | 7.41 (0 – 20.83) | 8.33 (0 – 20.83) | 7.64 (0 – 29.17) |

*Table 4: Mean percentage of errors in Experiment 3b across the three reward conditions and high and low probability sides. Errors are split into time-out errors, where participants did not make the manual response within the allocated time, and response errors. The range across participants is given in brackets.*

| Side | Condition | | | |
|---|---|---|---|---|
| | Probability | | Reward | |
| | Time-Out | Response Errors | Time-Out | Response Errors |
| High | 5.05 (0-9.17) | 4.86 (0.83-19.17) | 6.67 (1.39-16.67) | 6.85 (0-15.28) |
| Low | 18.87 (0-33.3) | 9.23 (0-25) | 10.34 (1.39-23.6) | 8.91 (0-27.78) |
| Equal | 8.3 (1.39-20.83) | 7.84 (0-18.06) | 7.68 (0-18.06) | 7.64 (0-16.67) |

*Table 5. The mean percentage of errors for Experiment 4 for both error types. The range across 18 participants is given in brackets.*

**Figures**

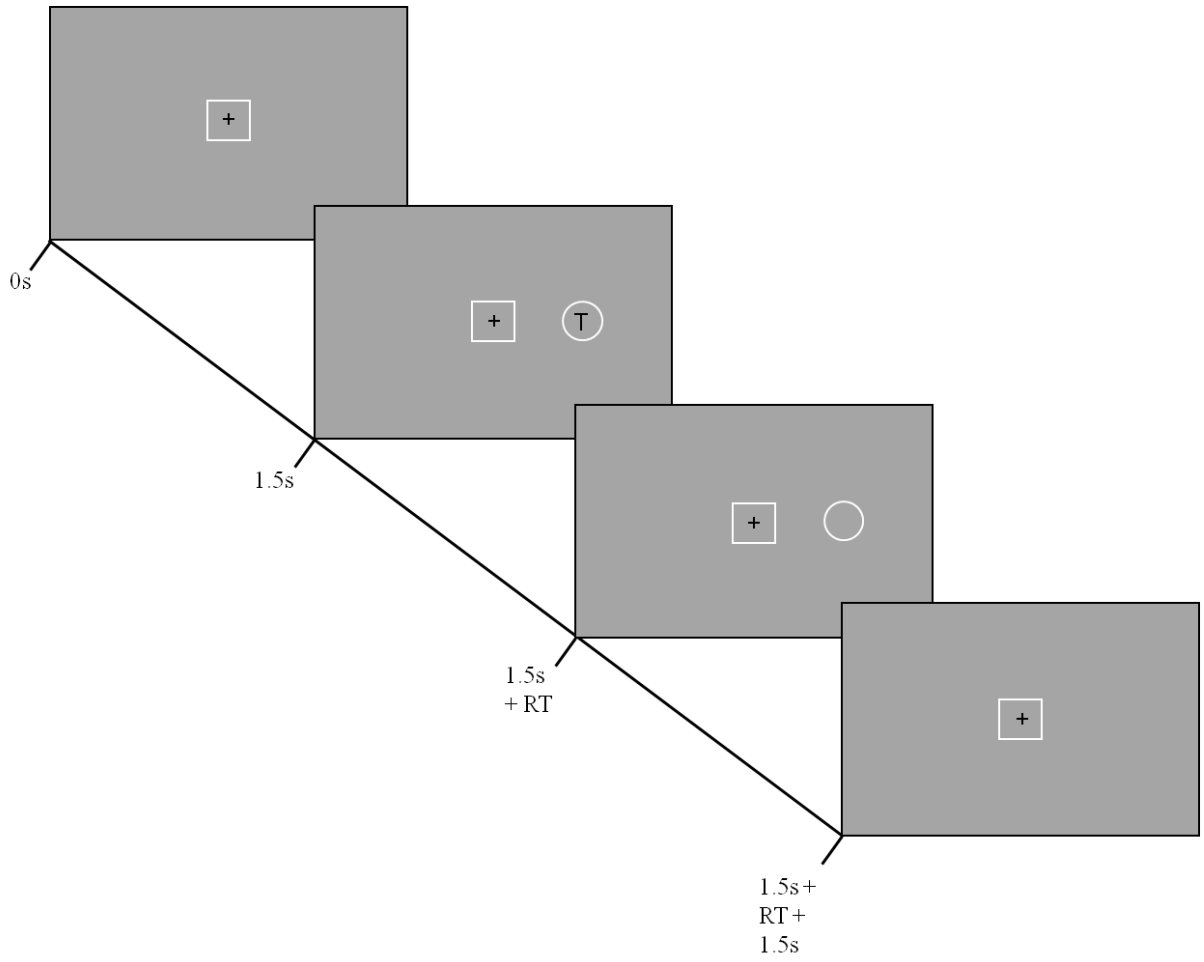


*Figure 1*

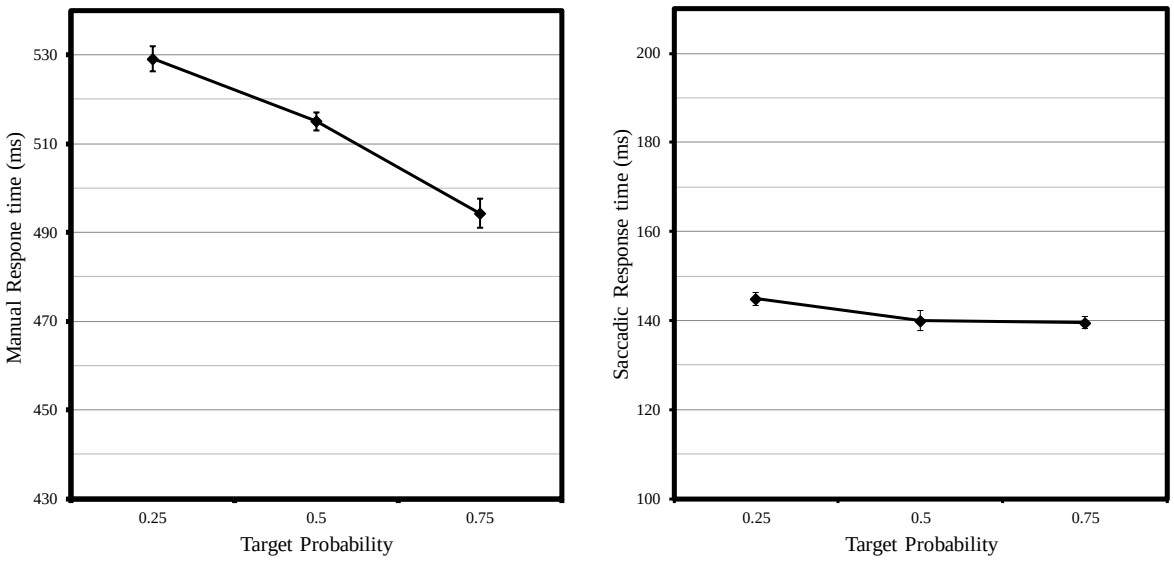


*Figure 2*

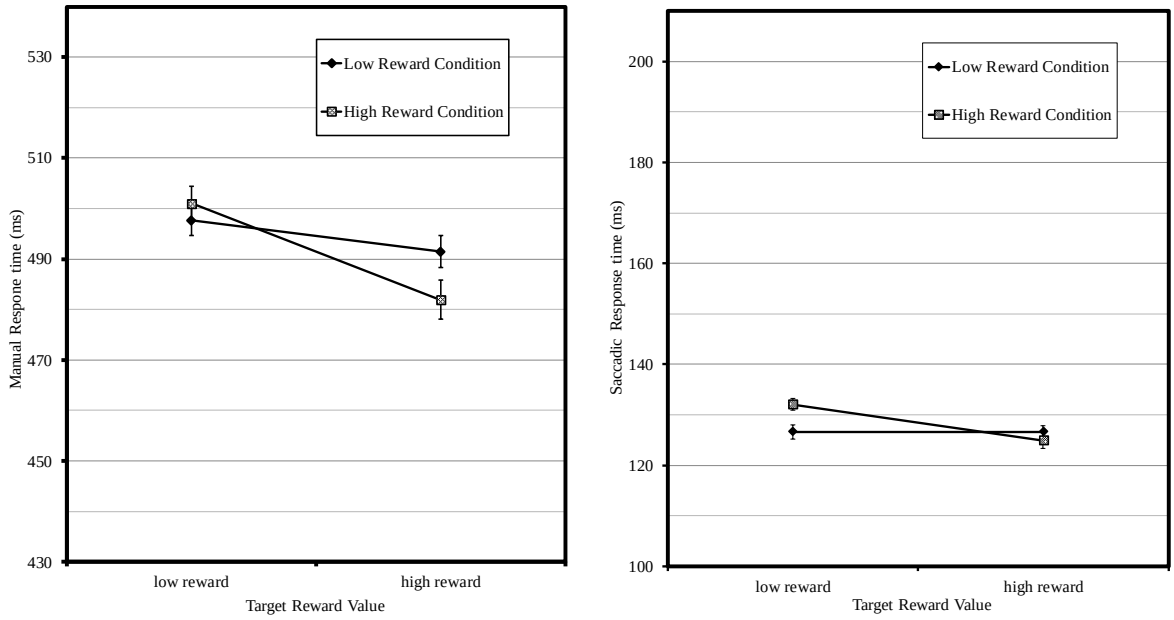


*Figure 3*

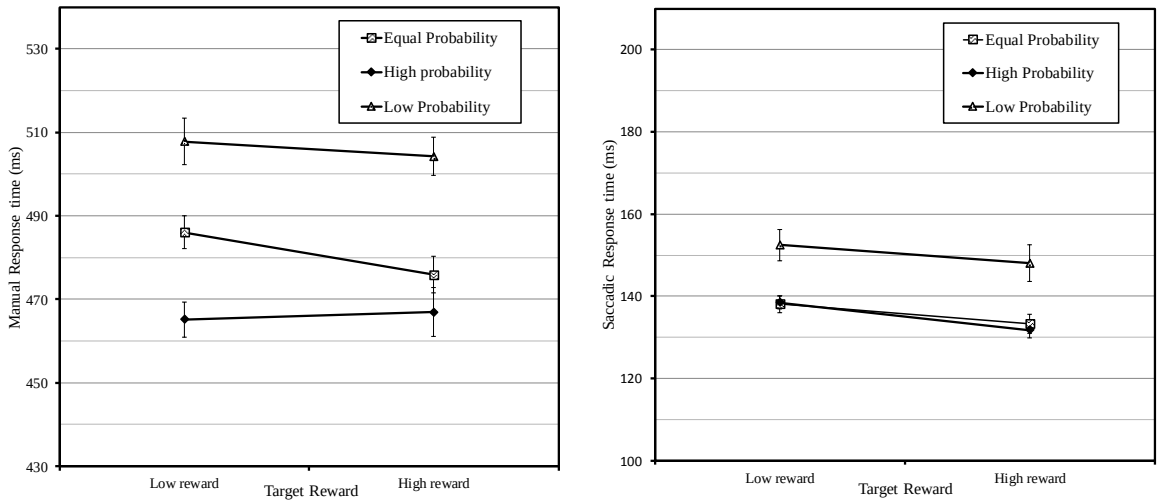
Equal Probability
High probability
Low Probability
Manual Response time (ms)
530
510
490
470
450
430
Low reward
Target Reward
High reward
Equal Probability
High Probability
Low Probability
Saccadic Response time (ms)
200
180
160
140
120
100
Low reward
Target Reward
High reward

*Figure 4*

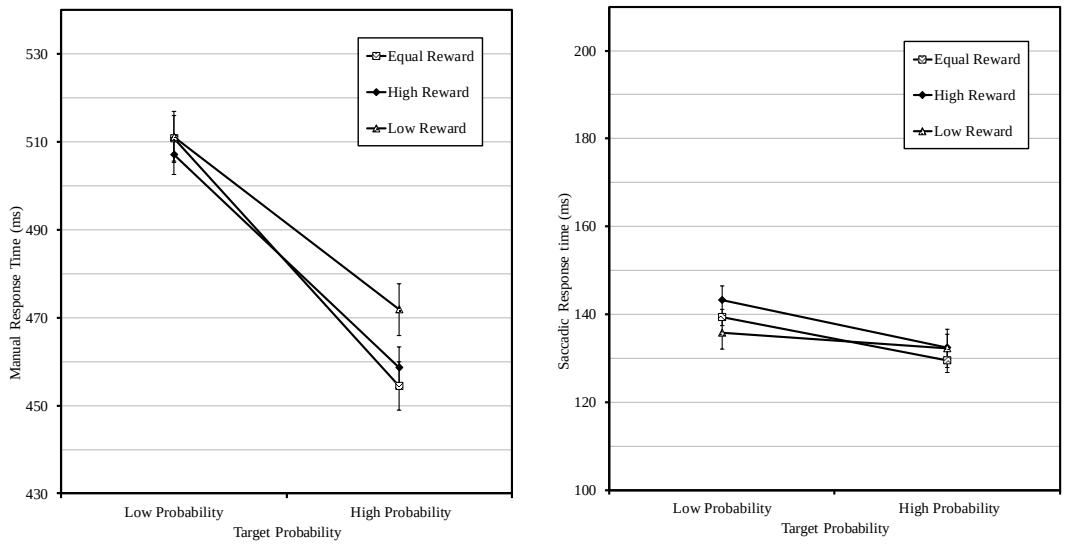


*Figure 5*

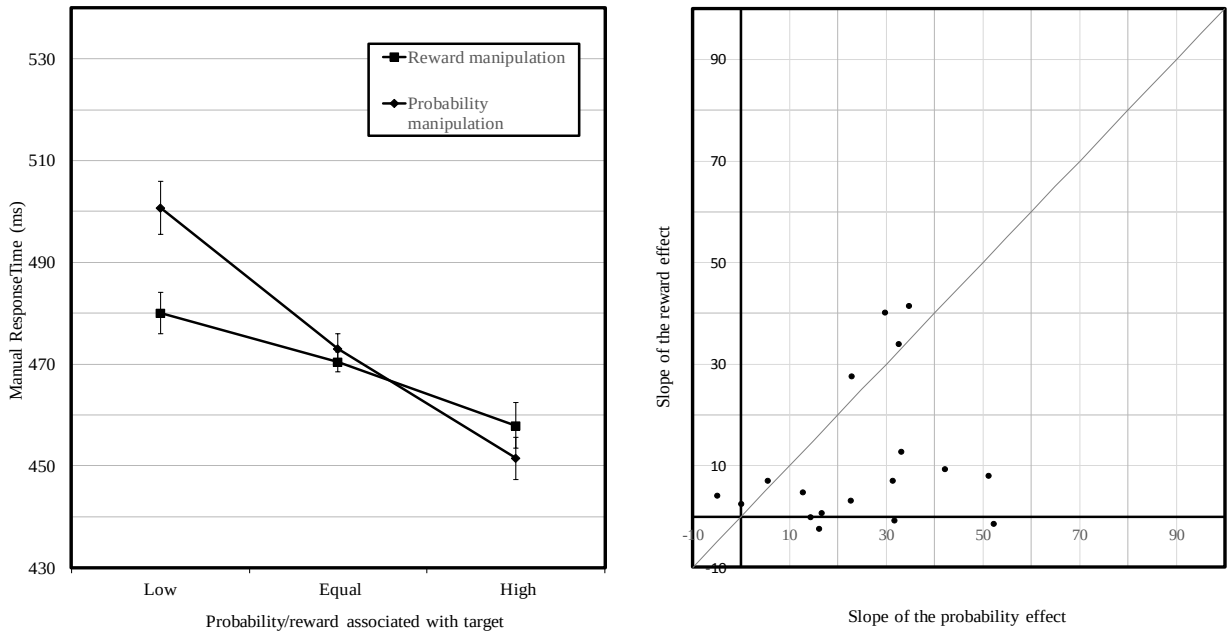


*Figure 6*

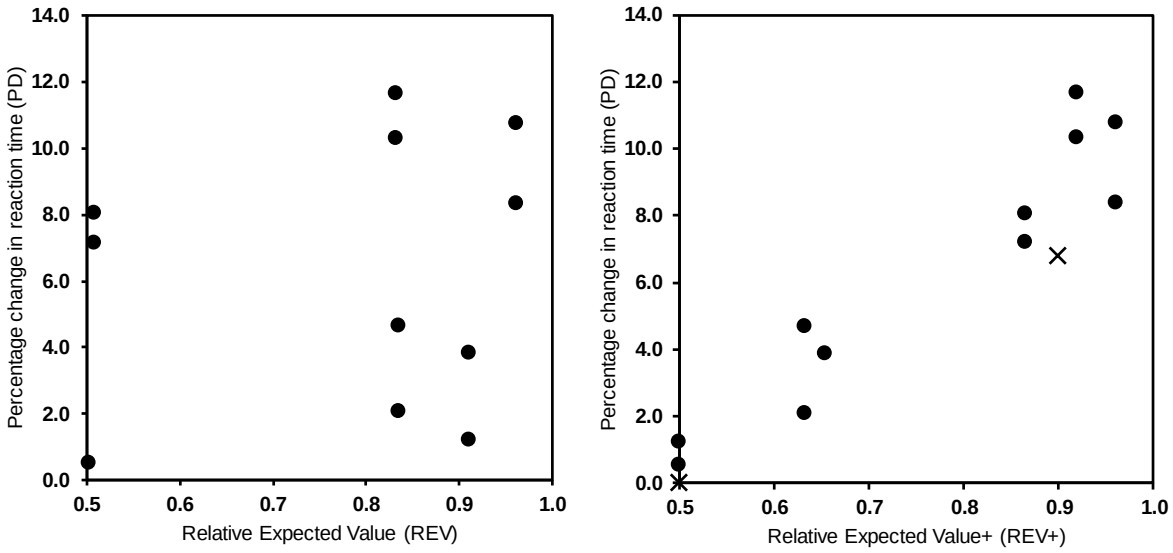


*Figure 7*

**Figure Captions**

*Figure 1. Diagram of the experimental procedure in the testing phase. The figure depicts an example trial where the target appears on the right-hand side.*

*Figure 2. The results from Experiment 1. Graph of mean of median manual response times (left hand plot) and saccadic response times (right hand plot) for Experiment 1. The three x-axis points denote the different probabilities associated with targets, with the data having been pooled from targets presented in the left and right hemispheres. The error bars are calculated within subjects and are the standard error of the mean.*

*Figure 3; Graph of mean of median manual response times (left hand side) and saccadic response times (right hand side) for Experiment 2. The two lines show the two different conditions (high reward magnitude and low reward magnitude) and the two x-axis points denote the reward associated with the targets (high and low reward). The within-subject error bars are the standard error of the mean.*

*Figure 4: Graph of the mean of median manual reaction times (left hand graph) and median saccadic reaction times (right hand graph) across participants for Experiment 3a. Each line represents a different probability, with the data split within blocks according to the probability associated with a particular target. The two x-axis points show the reward associated with the target. The within-subject error bars are the standard error of the mean.*

*Figure 5: Graph of the mean of median manual response times (left hand side) and saccadic response times (right hand side) for Experiment 3b. The three lines show the different reward conditions, which are grouped into low/high/equal reward rather than the conditions within the experiment. The x-axis gives the target probability. The within-subject error bars are the standard error of the mean.*

*Figure 6. Experiment 4. Left hand graph: manual response time. The two lines show the two different conditions (reward and probability) and the three x-axis points denote the level of reward or probability associated with the target. The within-subject error bars are the standard error of the mean. Right hand graph: Scatter plot of the slope of the probability effect against the slope of the reward effect across all participants.*

*Figure 7: Left-hand side: the relationship between Change in Reaction Time plotted against Relative Expected Value. The values plotted represent means from each condition across all 4 experiments. Right-hand side: the relationship between Change in Reaction Time plotted against Relative Expected Value+. The values plotted represent means from each condition across all 4 experiments.*